

CLAIMS

Having thus described our invention, what we claim as new and desire to secure by Letters Patent is as follows:

1 1. An automated method for setting up an instance of a natural language
2 conversational interface in a Web site comprising the steps of:
3 defining a hierarchy of topics into which individual documents or Web
4 pages can be classified;
5 generating a keyword index for those documents for an associated
6 search engine; and
7 for each node in the hierarchy, specifying a mechanism for associating
8 an input natural language (NL) query to the node.

1 2. The automated method for setting up an instance of a natural language
2 conversational interface in a Web site recited in claim 1, wherein the step of
3 generating a keyword index comprises the step of extracting sparse n-grams of
4 keywords for each group of pages in the topic hierarchy.

1 3. The automated method for setting up an instance of a natural language
2 conversational interface in a Web site recited in claim 1, further comprising
3 the step of optionally reviewing and editing the keyword index.

1 4. An automated method for setting up an instance of a natural language
2 interface in a Web site comprising the steps of:
3 automatically inducing a classification hierarchy by examining a
4 structure of the Web site;
5 creating index terms for leaf pages from sparse n-grams; and

6 creating rules for a classification engine from the sparse n-grams of
7 pages reachable from each node in a hierarchy of leaf pages, wherein each
8 node is a classification category and the rules associated with that category are
9 used to decide if a new input document or query reference the node.

1 5. The automated method for setting up an instance of a natural language
2 interface in a Web site recited in claim 4, wherein the step of creating rules for
3 a classification engine is performed automatically and further comprising the
4 optional step of manually editing the rules.